

Approximate Maximum Likelihood Source Localization from Range Measurements Through Convex Relaxation

Pınar Oğuz-Ekim, *Student Member, IEEE*, João Gomes, *Member, IEEE*,

João Xavier, *Member, IEEE*,

Marko Stošić, and Paulo Oliveira, *Senior Member, IEEE*

Abstract

This work considers the problem of locating a single source from noisy range measurements to a set of nodes in a wireless sensor network. We propose two new techniques that we designate as Source Localization with Nuclear Norm (SLNN) and Source Localization with ℓ_1 -norm (SL- ℓ_1), which extend to arbitrary real dimensions, including 3D, our prior work on 2D source localization formulated in the complex plane. Broadly, our approach is based on formulating a Maximum-Likelihood (ML) estimation problem for the source position, and then using convex relaxation techniques to obtain a semidefinite program (SDP) that can be globally and efficiently solved. SLNN directly approximates the Gaussian ML solution, and the relaxation is shown to be tighter than in other methods in the same class. We present an analysis of the convexity properties of the constraint set for the 2D complex version of SLNN (SLCP) to justify the observed tightness of the relaxation. In terms of global accuracy of localization, SLNN outperforms state-of-the-art optimization-based methods with either iterative or closed-form formulations. We propose the SL- ℓ_1 algorithm to address the Laplacian noise case, which models the presence of outliers in range measurements. We overcome the nondifferentiability of the Laplacian likelihood function by rewriting the ML problem as an exact weighted version of the Gaussian case, and compare two solution strategies. One of them is iterative, based on block coordinate descent, and uses SLNN as a subprocess block. The other, attaining only slightly worse performance, is noniterative and based on an SDP relaxation of the weighted ML problem.

The authors are with the Institute for Systems and Robotics – Instituto Superior Técnico (ISR/IST), Lisbon, Portugal. e-mails: {poguz.jpg,jxavier,mstosic,pjcro}@isr.ist.utl.pt.

This research was partially supported by Fundação para a Ciência e a Tecnologia (FCT) through ISR/IST plurianual funding with PIDDAC program funds, projects PTDC/EEA-TEL/71263/2006, PTDC/EEA-CRO/104243/2008, CMU-PT/SIA/0026/2009, and grant SFRH/BD/44771/2008.

I. INTRODUCTION

Locating a source from range measurements to a set of known reference points (anchors) is a classic problem in many engineering applications (e.g., radar, sonar, GPS), and has received a great deal of attention over the years. Recently, source localization from range measurements has been intensively examined in the context of wireless sensor networks (WSN), where ranges estimated from times of arrival, or from surrogates such as received signal strength, are somewhat unreliable due to the complexity of many WSN propagation environments (e.g., indoor settings with few unobstructed line-of-sight paths).

Spatial information *per se*, or as georeference to other sensor measurements, is crucial in WSN applications and warrants investigation into suitable localization algorithms. While many approaches to source localization based on classical triangulation or heuristic criteria can be found in the WSN literature [1], [2], our primary focus is on optimization-based methods formally derived from the likelihood function of observations, or related cost functions [3]–[8]. By doing so, we expect to take advantage of the optimality properties of maximum likelihood (ML) estimates to improve the robustness to perturbations in range measurements. We do not consider alternative/complementary measurements such as angles of arrival or time differences of arrival. We also assume cooperative localization scenarios where absolute ranges, as opposed to range differences to a reference sensor, are measured. These can be obtained either by synchronizing clocks and transmitting waveforms from the source at known times (beacon mode), or by initiating the transmission at a reference sensor and measuring the round trip time to the source and back (transponder mode).

Centralized ML algorithms for range-based source localization, which require the transmission of the full data set to a fusion node for processing, are proposed in [3], [8] under Gaussian noise and in [9] under Laplacian noise. These resort to semidefinite relaxation (SDR) to alleviate the problem of algorithmic convergence to undesirable local maxima of the likelihood function. A related alternative approach proposed in [4] solves a constrained least-squares (LS) problem using squared range (SR) measurements, subject to a quadratic constraint. This was shown to outperform, on average, the ML SDR approach of [8] whose relaxed solutions sometimes fail to produce meaningful source position vectors (rank one solutions). Another approach, proposed in [6], approximates the ML solution via second-order cone programming and a low-dimensional search.

Distributed algorithms for wireless sensor nodes, where the source location is iteratively determined through in-network processing at individual nodes and communication between neighbours, are also being very actively pursued [10]–[12]. These techniques, however, are not the focus of our work. We also note

that source localization can be viewed as a special instance of sensor network localization, where the positions of several sources/sensors are simultaneously determined from pairwise range measurements. Related algorithms based on semidefinite programming (SDP) have been developed for this class of problems [9], [13], and are relevant when there is significant uncertainty in anchor positions (see, e.g., [14] for a similar SDP approach to source localization with anchor uncertainty using range differences).

This paper develops an alternative to the source localization ML SDR method of [8]. We term this approach, originally proposed in [3] for 2D localization under Gaussian noise, Source Localization in the Complex Plane (SLCP). Our relaxation for the nonconvex and nonsmooth likelihood function is tighter than the one presented in [8], in the sense that the relaxed solution will more often have (near) rank-1, as required to obtain target coordinates by factorization. SLCP also outperforms the SR-LS method of [4], which iteratively solves a generalized trust-region subproblem and dispenses with factorization of rank-1 matrices, but undergoes some degradation with noisy measurements due to squaring of ranges in the cost function. The degradation of SR-LS becomes more severe in the presence of outliers [9], which commonly affect practical range measurement systems, e.g., when non-line-of-sight propagation occurs.

This paper expands upon the results of [3] in several ways:

- 1) We extend the framework of SLCP from 2D localization, which relied on a formulation where target and anchor coordinates were represented as complex numbers, to arbitrary (real) dimensions. We term the new SDR method Source Localization with Nuclear Norm (SLNN), as this norm arises naturally in the cost function of our relaxed optimization problem. Similarly to SLCP, SLNN offers a tight relaxation in most problem instances, and retains a performance advantage over SR-LS.
- 2) We provide a more complete analysis of the accuracy properties of SLCP, whose success in providing tight relaxations relies on certain parametrically defined sets in \mathbb{R}^2 being nearly convex. We discuss the convexity of the sets and how to trace the convex hull for any of them, from which convexity can be empirically assessed. For three-anchor scenarios we also examine a search-based alternative to SVD decomposition to extract the source coordinates from the solution of SLCP (a positive semidefinite matrix with near rank-1).
- 3) In [9] a modification of SLCP, termed $SL-\ell_1$, was introduced for ML source localization under Laplacian noise. This makes the algorithm robust to outlier measurements, a property that was observed in simulation even for non-Laplacian range errors. In this paper we provide a conceptually similar extension for source localization beyond 2D, consisting of a reformulation of the nondifferentiable log-likelihood function for Laplacian noise as a reweighted version of the Gaussian log-likelihood. We propose both single convex formulations and a simpler iterative optimization

algorithm, which repeatedly solves weighted SLNN problems followed by weight refinement.

The organization of the paper is as follows. In Section II, we formulate the ML source location problem under Gaussian or Laplacian noise. In Section III we derive the SLCP algorithm for 2D localization, we analyze the geometry of the associated optimization problem and the tightness of the relaxation (Subsection III-A), and we propose criteria for factorizing the SDR solution to recover the source coordinates (Subsection III-B). In Section IV we derive the SLNN algorithm, which extends SLCP to 3 (and higher) dimensions, and propose an iterative version of this algorithm that can handle Laplacian noise (Section IV-A). Section V illustrates the performances of the algorithms in simulation. Finally, conclusions are drawn in Section VI.

Throughout, both scalars and individual position vectors are represented by lowercase letters. Other vectors and matrices are denoted by boldface lowercase and uppercase letters, respectively. Individual components of matrix \mathbf{X} are written as x_{ij} and those of vector \mathbf{x} as x_i (the same notation would be used for a hypothetical position vector x_i , but the distinction between both should be clear from context). The superscript T (H) denotes the transpose (hermitian) of the given real (complex) vector or matrix, $\langle \cdot, \cdot \rangle$ denotes the inner product of two vectors, and $\text{tr}(\cdot)$ denotes the trace of a matrix. For symmetric matrix \mathbf{X} , $\mathbf{X} \succeq 0$ means that \mathbf{X} is positive semidefinite. We denote the Frobenius norm of matrix \mathbf{X} as $\|\mathbf{X}\|_F = \sqrt{\text{tr}(\mathbf{X}^H \mathbf{X})}$ and its nuclear norm as $\|\mathbf{X}\|_N = \text{tr}((\mathbf{X}^H \mathbf{X})^{\frac{1}{2}})$. Below, \mathbf{I}_m is the $m \times m$ identity matrix and $\mathbf{1}_m$ is the vector of m ones. The convex hull of set \mathcal{S} is denoted by $\text{co}(\mathcal{S})$.

II. PROBLEM FORMULATION

Let $x \in \mathbb{R}^n$ be the unknown source position, $a_i \in \mathbb{R}^n$, $i = 1, \dots, m$ be known sensor positions (anchors), and $r_i = \|x - a_i\| + w_i$ be the measured range between the source and the i -th anchor, where w_i denotes a noise term with standard deviation σ . Under i.i.d. Gaussian or Laplacian noise maximizing the likelihood of observations for the source localization problem is equivalent to

$$\underset{x}{\text{minimize}} \quad \sum_{i=1}^m \left| \|x - a_i\|^p - r_i^p \right|^q. \quad (1)$$

We will derive the SLCP/SLNN algorithms to (approximately) solve (1) under Gaussian noise ($p = 1$, $q = 2$), whereas SL- ℓ_1 will solve it under Laplacian noise ($p = 1$, $q = 1$). The case ($p = 2$, $q = 2$) is also of interest and corresponds to the cost function used in the SR-LS algorithm of [4], which is used to benchmark our algorithms. Note that the cost function for SR-LS is not a likelihood function, and it arises out of mathematical convenience.

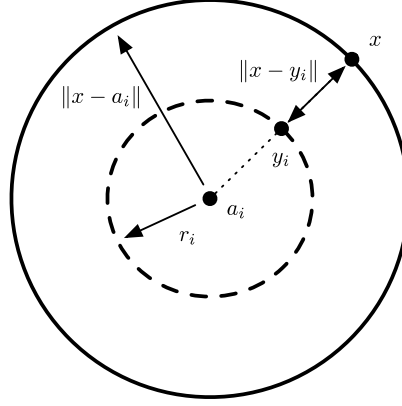


Fig. 1: Geometrical interpretation of terms in the source localization cost function (1) for $p = 1$, $q = 2$.

The main difficulties of solving (1) lie in the fact that this cost function is, in general, nonconvex and multimodal. For $q = 1$ it is also nondifferentiable, which poses additional challenges. We address the nonconvexity and multimodality of the cost function in Sections III and IV by developing convex relaxations that turn out to be tight in most problem instances, thus providing a very good approximation to the true source location. If necessary, the source coordinates can be further refined by iteratively minimizing (1) starting from the relaxed solution. Refer to [9] for one such iterative refinement approach based on the Majorization-Minimization algorithm. We address the nondifferentiability of (1) for $q = 1$ in Section IV-A by rewriting it as a weighted version of the case $q = 2$, where the weights themselves become optimization variables.

III. SOURCE LOCALIZATION IN 2D: SLCP

For $p = 1$, $q = 2$ we view each term in (1) as the squared distance between two circles centered on a_i , one with radius $\|x - a_i\|$, and the other with radius r_i (see Figure 1). This term can be replaced by the squared norm of the difference between the position vector x and its closest point on the circle $\{y \in \mathbb{R}^2 : \|y - a_i\| = r_i\}$, which we denote by y_i . Problem (1) can then be equivalently expressed as (a formal proof of equivalence is provided in [9])

$$\begin{aligned}
 & \underset{x, y_i}{\text{minimize}} && \sum_{i=1}^m \|x - y_i\|^2 \\
 & \text{subject to} && \|y_i - a_i\| = r_i \quad i = 1, \dots, m.
 \end{aligned} \tag{2}$$

If we fix y_i , the solution of (2) with respect to x is an unconstrained optimization problem whose solution is readily obtained as the center of mass of the constellation $x = \frac{1}{m} \sum_{i=1}^m y_i$. Moreover, in 2D

the constraints of (2) can be compactly described in the complex plane, yielding

$$\begin{aligned} & \text{minimize} \quad \left\| \frac{1}{m} \mathbf{1}_m \mathbf{1}_m^T \mathbf{y} - \mathbf{y} \right\|^2 \\ & \mathbf{y}, \boldsymbol{\theta} \\ & \text{subject to} \quad \mathbf{y} = \mathbf{a} + \mathbf{R}\boldsymbol{\theta}, \end{aligned} \quad (3)$$

where $\mathbf{a} = [a_1 \ \dots \ a_m]^T \in \mathbb{C}^m$ holds the anchor coordinates, expressed as complex numbers, $\mathbf{R} = \text{diag}(r_1, \dots, r_m) \in \mathbb{R}^{m \times m}$, and $\boldsymbol{\theta} = [e^{j\phi_1} \ \dots \ e^{j\phi_m}]^T \in \mathbb{C}^m$. The complex representation makes it simple to impose unit magnitude constraints on the elements of $\boldsymbol{\theta}$, and later relax them to obtain an SDR. Expanding the objective function and deleting constant terms yields the quadratic constrained problem

$$\begin{aligned} & \text{minimize} \quad 2 \text{Re}(\mathbf{c}^H \boldsymbol{\theta}) - \frac{1}{m} \boldsymbol{\theta}^H \mathbf{r} \mathbf{r}^T \boldsymbol{\theta} \\ & \boldsymbol{\theta} \\ & \text{subject to} \quad |\theta_i| = 1, \end{aligned} \quad (4)$$

where $\mathbf{r} = \mathbf{R} \mathbf{1}_m$ and $\mathbf{c} = \mathbf{R}(\mathbf{I}_m - \frac{1}{m} \mathbf{1}_m \mathbf{1}_m^T) \mathbf{a}$.

To proceed we now wish to replace $\text{Re}(\mathbf{c}^H \boldsymbol{\theta})$ in (4) with $-|\mathbf{c}^H \boldsymbol{\theta}|$, which is readily written as a function of a quadratic form in $\boldsymbol{\theta}$ and then relaxed in the same way as the second term in the objective function. To this end, first note that if $\boldsymbol{\theta}$ is replaced with $\boldsymbol{\theta} e^{j\gamma}$ neither the second term in the objective function of (4) nor the constraints change for any angle γ . By proper choice of γ the complex number $\mathbf{c}^H \boldsymbol{\theta}$ may be rotated to the (negative) real axis for any feasible $\boldsymbol{\theta}$, such that $\text{Re}(\mathbf{c}^H \boldsymbol{\theta} e^{j\gamma}) = -|\mathbf{c}^H \boldsymbol{\theta}|$, thus reducing the value of the objective function relative to other values of γ . This implies that any optimal solution of (4) will satisfy $\text{Re}(\mathbf{c}^H \boldsymbol{\theta}) = -|\mathbf{c}^H \boldsymbol{\theta}|$, which justifies replacing $\text{Re}(\cdot)$ with $-|\cdot|$ in the cost function. It should be kept in mind, however, that once a solution $\boldsymbol{\theta}$ to the modified optimization problem is obtained it should be rotated to obtain the actual vector of phases $\boldsymbol{\theta} e^{j\gamma}$ such that $\text{Re}(\mathbf{c}^H \boldsymbol{\theta} e^{j\gamma}) = -|\mathbf{c}^H \boldsymbol{\theta}|$.

Now the modified problem is equivalently written as

$$\begin{aligned} & \text{maximize} \quad 2\sqrt{\text{tr}(\mathbf{c} \mathbf{c}^H \boldsymbol{\theta} \boldsymbol{\theta}^H)} + \frac{1}{m} \text{tr}(\mathbf{r} \mathbf{r}^T \boldsymbol{\theta} \boldsymbol{\theta}^H) \\ & \boldsymbol{\theta} \\ & \text{subject to} \quad |\theta_i| = 1, \end{aligned} \quad (5)$$

and following standard manipulations we introduce the new variable $\Phi = \boldsymbol{\theta} \boldsymbol{\theta}^H$ and an associated (nonconvex) constraint $\text{rank}(\Phi) = 1$. Finally, a SDR formulation of SLCP is obtained by introducing the hypograph variable t such that $0 \leq t \leq 2\sqrt{\text{tr}(\mathbf{c} \mathbf{c}^H \Phi)}$ and dropping the rank constraint

$$\begin{aligned} & \text{maximize} \quad t + \frac{1}{m} \text{tr}(\mathbf{r} \mathbf{r}^T \Phi) \\ & \Phi, t \\ & \text{subject to} \quad \Phi \succeq 0, \ \phi_{ii} = 1, \ 4\mathbf{c}^H \Phi \mathbf{c} \geq t^2. \end{aligned} \quad (6)$$

TABLE I: Summary of the SLCP algorithm

- 1) Given the anchor positions and range measurements, solve the SDR (6)
- 2) Compute a rank-1 approximation of the SDR solution as $\Phi \approx \theta\theta^H$
- 3) Compute a rotation angle γ such that $\text{Re}(\mathbf{c}^H \theta e^{j\gamma}) = -|\mathbf{c}^H \theta|$ in (4)
- 4) Obtain the vector of circle projections $\mathbf{y} = \mathbf{a} + \mathbf{R}\theta e^{j\gamma}$
- 5) Estimate the source position as the centroid $x = \frac{1}{m} \mathbf{1}_m^T \mathbf{y}$

Remark that the solution of (6) is a positive semidefinite matrix, which should have a clearly dominant eigenvalue in problem instances where the SDR is an accurate approximation to the initial problem (2). In such cases $\Phi \approx \lambda_1 \mathbf{u}_1 \mathbf{u}_1^H$, where λ_1 is the highest eigenvalue of Φ and \mathbf{u}_1 the corresponding eigenvector, and the vector of complex phases is estimated as $\theta = \sqrt{\lambda_1} \mathbf{u}_1$ [15]. An alternative approach for computing θ is examined in Section III-B. Table I summarizes the SLCP algorithm.

A. Tightness and Geometry of the Constraint Set in SLCP

The source localization problem prior to relaxation (5) can be written as

$$\begin{aligned}
 & \text{maximize} && 2\sqrt{u} + \frac{1}{m}v \\
 & && u, v \\
 & \text{subject to} && (u, v) \in \mathcal{S},
 \end{aligned} \tag{7}$$

where

$$\mathcal{S} = \{(|\mathbf{c}^H \theta|^2, |\mathbf{r}^T \theta|^2) : \theta \in \mathbb{C}^m, |\theta_i| = 1\}. \tag{8}$$

The objective function in (7) is concave with respect to u and v , and the optimization problem would be convex if the set \mathcal{S} , over which this function should be maximized, were convex. Then, the SDR used in SLCP (6) would always find a rank-1 solution Φ , from which the vector of phases θ would readily follow by factorization. In practice it was found that, even for a moderate number of anchors, the set \mathcal{S} is likely to have the required shape along part of its border, as discussed below, so that the SDR solution has indeed rank-1. We now examine some of the properties of \mathcal{S} and the optimal solution.

Given the separable form of the cost function (7) it is clear that, for fixed v , it can be maximized by choosing u as large as possible within \mathcal{S} , and vice-versa. This implies the following property for the optimal points of (7):

Property 1. *The optimal points of (7) lie on the “upper right” boundary of set \mathcal{S} , i.e., optimal points of (7) are maximal elements of \mathcal{S} with respect to the standard cone \mathbb{R}_+^2 [16].*

Regarding the convexity properties of \mathcal{S} , recall that the cost function of (7) was designed to be invariant to rotations of $\boldsymbol{\theta}$ so that, without loss of generality, the first element may be taken as unity. For $m = 2$ anchors and $\theta_1 = 1$, $\theta_2 = e^{j\phi}$ we then have

$$u = |c_1^* + c_2^* e^{j\phi}|^2 = |c_1|^2 + |c_2|^2 + 2|c_1||c_2| \cos(\phi + \alpha) \quad (9)$$

$$v = |r_1 + r_2 e^{j\phi}|^2 = r_1^2 + r_2^2 + 2r_1 r_2 \cos \phi, \quad (10)$$

where $\alpha = \angle c_1 - \angle c_2$. Set \mathcal{S} is an ellipse centered on $(|c_1|^2 + |c_2|^2, r_1^2 + r_2^2)$, therefore clearly nonconvex. Given the definitions of \mathbf{c} and \mathbf{r} in (4), for $m > 2$ anchors it is always possible to zero out elements $3, \dots, m$ in these vectors if $r_3 = \dots = r_m = 0$ in the diagonal of \mathbf{R} , thus reverting to the case $m = 2$. In summary:

Property 2. *Depending on the specific range measurements, set \mathcal{S} may be nonconvex for any number of anchors.*

In spite of the lack of convexity guarantees for \mathcal{S} , our simulation results suggest that for $m \geq 3$ anchors and typical range measurements this set usually does have a convex-like shape. Even when \mathcal{S} is not convex all that is required for our SDR to provide a rank-1 solution is “local convexity” along the “upper right” boundary of \mathcal{S} where the optimal point of (7) is known to be located. More formally, we require that the intersection of \mathcal{S} with any supporting hyperplane defined by a normal direction with nonnegative components be a compact subset (a single point or a line segment) [16]. Figure 2 depicts some examples of \mathcal{S} for different numbers of anchors and randomly generated \mathbf{c} , \mathbf{r} . Our practical test for local (non)convexity of \mathcal{S} consists of tracing multiple supporting hyperplanes with nonnegative normal elements, and assessing whether any of them intersect \mathcal{S} at two well-separated points. We build supporting hyperplanes not on \mathcal{S} directly, which is a hard problem, but on the related relaxed convex set

$$\mathcal{T} = \{(\text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\Phi}), \text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\Phi})) : \boldsymbol{\Phi} \in \mathbb{C}^{m \times m}, \boldsymbol{\Phi} \succeq 0, \phi_{ii} = 1\}. \quad (11)$$

Specifically, for a supporting hyperplane with normal $(\cos \beta, \sin \beta)$, $0 \leq \beta \leq \frac{\pi}{2}$, we determine an intersection point with \mathcal{T} by solving the convex optimization problem

$$\begin{aligned} & \underset{\boldsymbol{\Phi}}{\text{maximize}} && \langle (\cos \beta, \sin \beta), (\text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\Phi}), \text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\Phi})) \rangle \\ & && \boldsymbol{\Phi} \\ & \text{subject to} && \boldsymbol{\Phi} \succeq 0, \phi_{ii} = 1, \end{aligned} \quad (12)$$

and setting the intersection point as $u = \text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\Phi})$, $v = \text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\Phi})$. This procedure is justified by the following result, proved in Appendix A.

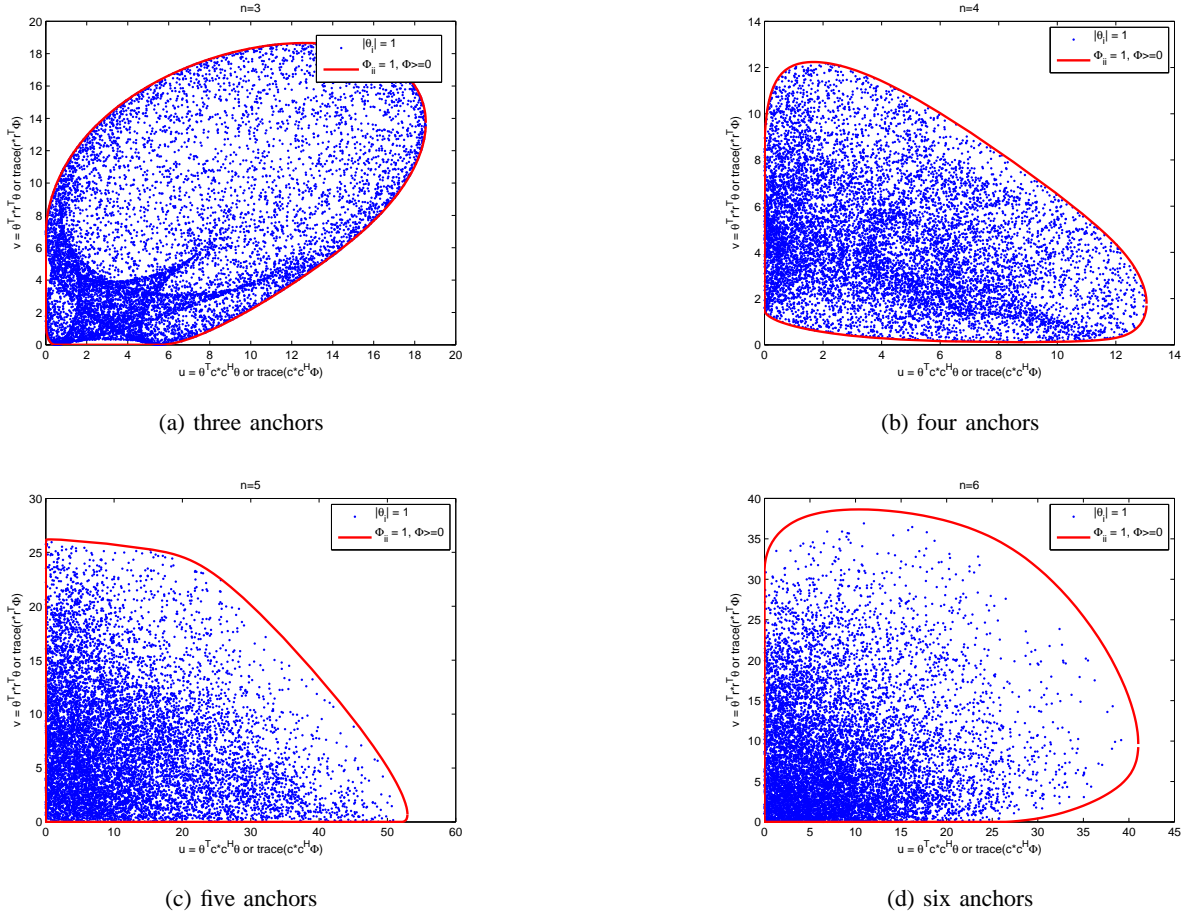


Fig. 2: The constraint set \mathcal{S} for randomly generated θ satisfying $|\theta_i| = 1$ and different numbers of randomly placed anchors. For each set the hypothesized convex hull, computed by relaxation of \mathcal{S} , is also depicted.

Lemma 1. *For $m \leq 3$ anchors the sets \mathcal{S} and \mathcal{T} have the same set of supporting hyperplanes with nonnegative normal elements. Equivalently, in the relevant portion of its boundary \mathcal{T} coincides with the convex hull of \mathcal{S} .*

Although we only prove this result up to $m = 3$, the empirical evidence suggests that it is also valid for higher m , at least up to some maximum order (see Figure 2). We leave this as a conjecture and apply the procedure for $m > 3$ as well, noting, however, that the case $m = 3$ has major practical significance as the minimum number of anchors that are necessary to recover a general 2D source position based on range measurements. We also conjecture that \mathcal{T} is actually the convex hull of \mathcal{S} , so (12) may be used to trace

the full boundary of $\text{co}(\mathcal{S})$, and not just the portion where the supporting hyperplanes have nonnegative normal elements. This assumption is not required for our analysis, but was used for generating the set boundaries shown in Figure 2.

B. Factorization of the SDR Solution

The solution of the relaxed SLCP optimization problem (6) is a positive semidefinite matrix, Φ , from which the vector of complex exponentials θ is calculated by rank-1 factorization. The latter is needed to form the vector of circle projections $\mathbf{y} = \mathbf{a} + \mathbf{R}\theta$ (see (3)) and, ultimately, the source position vector as the centroid $x = \frac{1}{m}\mathbf{1}_m^T \mathbf{y}$. The rank-1 factorization method advocated at the end of Section III is truncation of the eigenvalue decomposition of Φ at the highest eigenvalue. Here we examine a more exact search-based alternative for the practically relevant case of $m = 3$ anchors, which will also be useful to assess the accuracy of the factorization based on eigenvalue truncation.

For a given positive semidefinite matrix $\Phi \in \mathbf{C}^{m \times m}$ we wish to find vector $\theta \in \mathbf{C}^m$ satisfying

$$\begin{aligned} & \text{minimize} \quad \|\Phi - \theta\theta^H\|_F^2 \\ & \quad \theta \\ & \text{subject to} \quad |\theta_i| = 1. \end{aligned} \tag{13}$$

The objective function in (13) is expanded as

$$\|\Phi - \theta\theta^H\|_F^2 = \text{tr}((\Phi - \theta\theta^H)^H(\Phi - \theta\theta^H)) = \|\Phi\|_F^2 + \underbrace{\|\theta\|^4}_{m^2} - \text{tr}(\Phi^H \theta\theta^H) - \text{tr}(\theta\theta^H \Phi). \tag{14}$$

Ignoring constant terms the optimization problem is equivalently reformulated as

$$\begin{aligned} & \text{maximize} \quad \theta^H \Phi \theta \\ & \quad \theta \\ & \text{subject to} \quad |\theta_i| = 1. \end{aligned} \tag{15}$$

The cost function of (15) is insensitive to a global rotation of all elements of θ by a common factor, hence for $m = 3$ anchors θ can be written as $\theta = \begin{bmatrix} 1 & e^{j\alpha} & e^{j(\alpha+\delta)} \end{bmatrix}^T$ and (15) becomes

$$\begin{aligned} & \text{maximize} \quad \text{Re}(\phi_{12}e^{j\alpha} + \phi_{23}e^{j\delta} + \phi_{13}e^{j(\alpha+\delta)}). \\ & \quad \alpha, \delta \end{aligned} \tag{16}$$

For fixed α the maximum is attained for $\delta = -\angle(\phi_{23} + \phi_{13}e^{j\alpha})$, yielding for (16)

$$\begin{aligned} & \text{maximize} \quad \text{Re}(\phi_{12}e^{j\alpha}) + |\phi_{23} + \phi_{13}e^{j\alpha}|. \\ & \quad \alpha \end{aligned} \tag{17}$$

The solution to (17) is found by searching for the maximum value over the interval $[0, 2\pi)$.

Referring to the definitions of the 2D sets \mathcal{S} in (8) and \mathcal{T} in (11), similar criteria to the above were considered for finding θ such that the induced point in \mathcal{S} is closest in Euclidean norm to the one induced by Φ in \mathcal{T} . However, the many-to-one nature of the mapping of θ onto points in \mathcal{S} makes this formulation intrinsically ambiguous.

IV. SOURCE LOCALIZATION IN HIGHER DIMENSIONS: SLNN

To extend the approach used in SLCP to $n > 2$ dimensions, we write the circle/sphere equations in (2) using an equivalent parametric form with real coordinates

$$\begin{aligned} & \text{minimize} \quad \sum_{i=1}^m \|x - y_i\|^2 \\ & \quad x, y_i, u_i \\ & \text{subject to} \quad y_i = a_i + r_i u_i, \quad \|u_i\| = 1, \end{aligned} \tag{18}$$

where x, y_i, a_i and u_i are now vectors in \mathbb{R}^n , rather than complex scalars used in SLCP. In (18) $u_i \in \mathbb{R}^n$ is a unit-norm vector that plays the same role as the complex phase shift $e^{j\phi_i}$ in SLCP. Equivalently,

$$\begin{aligned} & \text{minimize} \quad \|\mathbf{1}_m x^T - \mathbf{Y}\|_F^2 \\ & \quad x, y_i, u_i \\ & \text{subject to} \quad \underbrace{\begin{bmatrix} y_1^T \\ \vdots \\ y_m^T \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}}_{\mathbf{A}} + \mathbf{R} \underbrace{\begin{bmatrix} u_1^T \\ \vdots \\ u_m^T \end{bmatrix}}_{\mathbf{U}}, \quad \|u_i\| = 1, \end{aligned} \tag{19}$$

where $\mathbf{R} = \text{diag}(r_1, \dots, r_m)$ as in (3). For fixed y_i, u_i (19) describes n uncoupled least-squares problems whose variables are the components of the source location vector x . The optimal solutions may be jointly written compactly as

$$x^T = (\mathbf{1}_m^T \mathbf{1}_m)^{-1} \mathbf{1}_m^T \mathbf{Y} = \frac{1}{m} \mathbf{1}_m^T \mathbf{Y}. \tag{20}$$

Replacing this back in (19) to eliminate variable x the objective function becomes $\|\mathbf{\Pi Y}\|_F^2 = \text{tr}(\mathbf{Y}^T \mathbf{\Pi Y})$, where $\mathbf{\Pi} = \mathbf{I}_m - \frac{1}{m} \mathbf{1}_m \mathbf{1}_m^T$ is a projection matrix (hence idempotent). Similarly to (3)–(4) we can now eliminate variable \mathbf{Y} and the first set of equality constraints, expanding its definition in the objective function and ignoring constant terms to obtain

$$\begin{aligned} & \text{minimize} \quad 2 \text{tr}(\mathbf{C}^T \mathbf{U}) - \frac{1}{m} \text{tr}(\mathbf{U}^T \mathbf{r} \mathbf{r}^T \mathbf{U}) \\ & \quad \mathbf{U} \\ & \text{subject to} \quad \|u_i\| = 1, \end{aligned} \tag{21}$$

where $\mathbf{C} = \mathbf{R} \mathbf{\Pi A}$ and, as in (4), $\mathbf{r} = \mathbf{R} \mathbf{1}_m$.

a) *Nuclear Norm Approximation:* As in the complex formulation we wish to rewrite the first term in the objective function of (21) in a form that is more amenable to SDR. In the optimization problem we thus replace \mathbf{U} with the product $\mathbf{U}\mathbf{V}$, where \mathbf{V} is an $n \times n$ orthogonal matrix such that $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I}_n$, yielding

$$\begin{aligned} & \text{minimize} \quad 2\text{tr}(\mathbf{C}^T\mathbf{U}\mathbf{V}) - \frac{1}{m}\text{tr}(\mathbf{V}^T\mathbf{U}^T\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{V}) \\ & \quad \mathbf{U}, \mathbf{V} \\ & \text{subject to} \quad \|u_i\| = 1, \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}_n. \end{aligned} \quad (22)$$

Note that, due to the orthogonality of \mathbf{V} , each line of $\mathbf{U}\mathbf{V}$ still has unit norm, so for any feasible \mathbf{U} in (21) $\mathbf{U}\mathbf{V}$ is also feasible. Regarding (22), \mathbf{V} may be interpreted as an inner optimization variable that, for each candidate \mathbf{U} , minimizes the value of the objective function. Noting that the second term in the objective function (22) does not depend on \mathbf{V} , as $\text{tr}(\mathbf{V}^T\mathbf{U}^T\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{V}) = \text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{V}\mathbf{V}^T\mathbf{U}^T) = \text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{U}^T)$, the inner optimization problem simply becomes

$$\begin{aligned} & \text{minimize} \quad \text{tr}(\mathbf{C}^T\mathbf{U}\mathbf{V}) = \langle \mathbf{V}, \mathbf{U}^T\mathbf{C} \rangle \\ & \quad \mathbf{V} \\ & \text{subject to} \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}_n. \end{aligned} \quad (23)$$

This involves the minimization of a linear function on the set of orthogonal matrices, which resembles the known problem of minimizing a linear function of a vector \mathbf{v} , say, $\langle \mathbf{v}, \mathbf{a} \rangle$, on the unit sphere $\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v} = 1$. Invoking the KKT conditions [16] the latter problem is readily seen to yield the optimal cost $-\|\mathbf{a}\|$, attained at the point on the sphere along vector $-\mathbf{a}$. One would therefore expect the solution of (23) to be $-\|\mathbf{C}^T\mathbf{U}\|$, involving some matrix norm of $\mathbf{C}^T\mathbf{U}$. In Appendix B it is shown that this is indeed the case, and that the appropriate norm to consider is the *nuclear norm*, defined for matrix \mathbf{X} as $\|\mathbf{X}\|_N = \text{tr}((\mathbf{X}^H\mathbf{X})^{\frac{1}{2}})$, and equaling the sum of its singular values [17]. The optimization problem (22) is therefore equivalently rewritten as

$$\begin{aligned} & \text{minimize} \quad -2\|\mathbf{C}^T\mathbf{U}\|_N - \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{U}^T) \\ & \quad \mathbf{U} \\ & \text{subject to} \quad \|u_i\| = 1, \end{aligned} \quad (24)$$

or

$$\begin{aligned} & \text{maximize} \quad 2\text{tr}((\mathbf{C}^T\mathbf{U}\mathbf{U}^T\mathbf{C})^{\frac{1}{2}}) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{U}^T) \\ & \quad \mathbf{U} \\ & \text{subject to} \quad \|u_i\| = 1. \end{aligned} \quad (25)$$

TABLE II: Summary of the SLNN algorithm

- 1) Given the anchor positions and range measurements, solve the SDR (27)
- 2) Compute a rank- n approximation of the SDR solution as $\mathbf{W} \approx \mathbf{U}\mathbf{U}^T$
- 3) Solve the inner optimization problem (23) to get the rotation matrix \mathbf{V}
- 4) Obtain the matrix of sphere projections as $\mathbf{Y} = \mathbf{A} + \mathbf{R}\mathbf{U}\mathbf{V}$
- 5) Estimate the source position as the centroid of the rows of \mathbf{Y} , $x = \frac{1}{m}\mathbf{Y}^T \mathbf{1}_m$

We now introduce the variable $\mathbf{W} = \mathbf{U}\mathbf{U}^T$ and ignore the associated nonconvex constraint $\text{rank}(\mathbf{W}) = n$ to obtain the SDR

$$\begin{aligned} & \underset{\mathbf{W}}{\text{maximize}} \quad 2 \text{tr}((\mathbf{C}^T \mathbf{W} \mathbf{C})^{\frac{1}{2}}) + \frac{1}{m} \text{tr}(\mathbf{r} \mathbf{r}^T \mathbf{W}) \\ & \text{subject to} \quad \mathbf{W} \succeq 0, \quad w_{ii} = 1. \end{aligned} \quad (26)$$

The objective function of (26) is the sum of a concave¹ function of \mathbf{W} with a linear term, and is therefore concave. The constraint set of (26) is convex, thus establishing that this is indeed a convex optimization problem. We express it in standard SDP form as

$$\begin{aligned} & \underset{\mathbf{W}, \mathbf{Z}}{\text{maximize}} \quad 2 \text{tr}(\mathbf{Z}) + \frac{1}{m} \text{tr}(\mathbf{r} \mathbf{r}^T \mathbf{W}) \\ & \text{subject to} \quad \mathbf{W} \succeq 0, \quad w_{ii} = 1, \quad \begin{bmatrix} \mathbf{C}^T \mathbf{W} \mathbf{C} & \mathbf{Z} \\ \mathbf{Z} & \mathbf{I}_n \end{bmatrix} \succeq 0, \quad \mathbf{Z} \succeq 0. \end{aligned} \quad (27)$$

The equivalence between (26) and (27) is proved in Appendix B.

Similarly to the complex 2D formulation, the solution of our SDR is a $m \times m$ matrix \mathbf{W} that should have approximately rank n when the relaxation is tight. The matrix \mathbf{U} of unit-norm vectors is obtained by SVD factorization of \mathbf{W} [15] and, after accounting for the inner rotation of \mathbf{U} , it is used to build the y_i and, ultimately, the source position vector x . Table II summarizes the SLNN algorithm.

A. Localization under Laplacian Noise: $SL\text{-}\ell_1$

When disturbances are Laplacian and i.i.d., thus heavier tailed than Gaussian, maximizing the likelihood amounts to solving (1) for $p = q = 1$,

$$\underset{x}{\text{minimize}} \quad \sum_{i=1}^m |||x - a_i|| - r_i|. \quad (28)$$

¹The first term is the composition of the linear map $\mathbf{X} = \mathbf{C}^T \mathbf{W} \mathbf{C}$ with $\text{tr}(\mathbf{X}^{\frac{1}{2}})$, which is known to be concave in \mathbf{X} [16].

The presence of $|\cdot|$ in each summation term of (28), rather than $(\cdot)^2$, de-emphasizes the contributions of measurements r_i corrupted by large noise values. The optimal point of (28) is thus less biased by these outlier measurements than the cost function (1) for the Gaussian case $p = 1$, $q = 2$. However, a major difficulty in solving (28) is the fact that the cost function is not differentiable, making it less amenable to the types of analytic manipulations that we use to develop SDR. The strategy that we adopt to circumvent this difficulty parallels the one used in [9] for 2D sources, and as a key ingredient involves squaring the cost function of (28) (which does not affect the location of extremal points), and then rewriting it as

$$\begin{aligned} & \text{minimize} \quad \sum_{i=1}^m \frac{(\|x - a_i\| - r_i)^2}{\lambda_i} \\ & \quad x, \boldsymbol{\lambda} \\ & \text{subject to} \quad \lambda_i > 0, \quad \mathbf{1}_m^T \boldsymbol{\lambda} = 1. \end{aligned} \quad (29)$$

The cost function is thus reduced to a weighted version of the more tractable Gaussian log-likelihood, where the real weighting coefficients λ_i become optimization variables themselves. See [9] for a proof of this result (also [18]). Now, the manipulations used earlier in Section IV for the development of SLNN can be replicated here to reformulate the problem as

$$\begin{aligned} & \text{minimize} \quad \sum_{i=1}^m \frac{\|x - y_i\|^2}{\lambda_i} \\ & \quad x, y_i, u_i, \boldsymbol{\lambda} \\ & \text{subject to} \quad y_i = a_i + r_i u_i, \quad \|u_i\| = 1, \quad \lambda_i > 0, \quad \mathbf{1}_m^T \boldsymbol{\lambda} = 1. \end{aligned} \quad (30)$$

For given y_i , u_i , and $\boldsymbol{\lambda}$, (30) has a least-squares cost function whose unconstrained optimal solution with respect to x is readily found in closed form from the first-order stationarity condition

$$\sum_{i=1}^m \frac{x - y_i}{\lambda_i} = 0, \quad x^* = \frac{\sum_{i=1}^m \frac{y_i}{\lambda_i}}{\sum_{i=1}^m \frac{1}{\lambda_i}}. \quad (31)$$

Substituting the optimal x in (30), and using matrix notation, the cost function becomes $\text{tr}(\mathbf{Y}^T \boldsymbol{\Xi} \mathbf{Y})$, where $\boldsymbol{\Xi}$ is the modified projector

$$\boldsymbol{\Xi} = \begin{bmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_m} \end{bmatrix} - \frac{1}{\sum_{i=1}^m \frac{1}{\lambda_i}} \begin{bmatrix} \frac{1}{\lambda_1} \\ \vdots \\ \frac{1}{\lambda_m} \end{bmatrix} \begin{bmatrix} \frac{1}{\lambda_1} & \dots & \frac{1}{\lambda_m} \end{bmatrix} = \boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1} \mathbf{1} (\mathbf{1}^T \boldsymbol{\Lambda}^{-1} \mathbf{1})^{-1} \mathbf{1}^T \boldsymbol{\Lambda}^{-1} \quad (32)$$

with $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_m)$.

1) *Alternating directions (SL- ℓ_1 AD)*: One possibility for iteratively solving (30) is to use block coordinate descent, alternating between minimizing the expression with respect to $\{x, y_i, u_i\}$ for fixed

λ and vice-versa. For fixed λ the problem is

$$\begin{aligned} & \text{minimize} \quad \text{tr}((\mathbf{A} + \mathbf{R}\mathbf{U})^T \mathbf{\Xi} (\mathbf{A} + \mathbf{R}\mathbf{U})) \\ & \quad \mathbf{U} \\ & \text{subject to} \quad \|u_i\| = 1, \end{aligned} \tag{33}$$

which differs from the SLNN formulation only in the projector matrix $\mathbf{\Xi}$. It can therefore be similarly manipulated into a relaxed form that parallels (27)

$$\begin{aligned} & \text{maximize} \quad 2 \text{tr}(\mathbf{Z}) + \frac{1}{\kappa} \text{tr}(\mathbf{r}\mathbf{r}^T \mathbf{W}) \\ & \quad \mathbf{W}, \mathbf{Z} \\ & \text{subject to} \quad \mathbf{W} \succeq 0, \quad w_{ii} = 1, \quad \begin{bmatrix} \mathbf{C}^T \mathbf{W} \mathbf{C} & \mathbf{Z} \\ \mathbf{Z} & \mathbf{I}_n \end{bmatrix} \succeq 0, \quad \mathbf{Z} \succeq 0, \end{aligned} \tag{34}$$

with $\mathbf{C} = \mathbf{R}\mathbf{\Xi}\mathbf{A}$, $\mathbf{r} = \mathbf{R} \begin{bmatrix} \frac{1}{\lambda_1} & \dots & \frac{1}{\lambda_m} \end{bmatrix}^T$, and $\kappa = \sum_{i=1}^m \frac{1}{\lambda_i}$. For the converse block coordinate descent step with fixed $\{x, y_i, u_i\}$ the problem is

$$\begin{aligned} & \text{minimize} \quad \sum_{i=1}^m \frac{K_i^2}{\lambda_i} \\ & \quad \lambda \\ & \text{subject to} \quad \lambda_i > 0, \quad \mathbf{1}_m^T \lambda = 1, \end{aligned} \tag{35}$$

where $K_i \triangleq |||x - a_i|| - r_i| = |||x - y_i||$ are constant in this subproblem. The solution, readily obtained from the first-order KKT conditions, is given by [9]

$$\lambda_i^* = \frac{K_i}{\sum_{i=1}^m K_i} = \frac{|||x - a_i|| - r_i|}{\sum_{i=1}^m |||x - a_i|| - r_i|}, \tag{36}$$

yielding the desired ℓ_1 -type cost function $\sum_{i=1}^m \frac{K_i^2}{\lambda_i} = (\sum_{i=1}^m K_i)^2$. We sequentially perform the iterations (34) and (36), starting with $\lambda = \frac{1}{m} \mathbf{1}_m$, until $\|x^{k+1} - x^k\|$ is within some prescribed tolerance ε . In our simulated scenarios on the order of 3–10 iterations are needed for $\varepsilon = 10^{-2}$. This method is denoted by SL- ℓ_1 AD.

2) *Non-iterative formulation (SL- ℓ_1 MD)*: For a non-iterative solution of (30) we start from the equivalent formulation (33), with $\lambda_1, \dots, \lambda_m$ included as optimization variables through the weighting matrix $\mathbf{\Xi}$, and introduce an epigraph variable t_i for each term contributing to $\text{tr}(\cdot)$ in the cost function

$$\begin{aligned} & \text{minimize} \quad \mathbf{t} \mathbf{1}_n \\ & \quad \mathbf{U}, \lambda, \mathbf{t} \\ & \text{subject to} \quad \mathbf{e}_i^T (\mathbf{A} + \mathbf{R}\mathbf{U})^T \mathbf{\Xi} (\mathbf{A} + \mathbf{R}\mathbf{U}) \mathbf{e}_i \leq t_i \\ & \quad \|u_i\| = 1, \quad \lambda_i > 0, \quad \mathbf{1}_m^T \lambda = 1, \end{aligned} \tag{37}$$

where $\mathbf{t} = [t_1 \ \dots \ t_n]$ and \mathbf{e}_i is the standard coordinate vector with 1 in the i -th position and zeros elsewhere. As in [9] we invoke the matrix inversion lemma to express (32) as the limiting case of (positive semidefinite) $\Xi = \lim_{\sigma \rightarrow \infty} (\mathbf{A} + \sigma \mathbf{1}_m \mathbf{1}_m^T)^{-1}$, which is more amenable to analytic manipulations in optimization problems. In practice we take σ as a sufficiently large constant. Using Schur complements the inequality constraint in (37) may be successively written as

$$\begin{bmatrix} t_i & \mathbf{e}_i^T (\mathbf{A} + \mathbf{R}\mathbf{U})^T \\ (\mathbf{A} + \mathbf{R}\mathbf{U})\mathbf{e}_i & \Xi^{-1} \end{bmatrix} \succeq 0 \quad (38)$$

$$t_i(\mathbf{A} + \sigma \mathbf{1}_m \mathbf{1}_m^T) - (\mathbf{A} + \mathbf{R}\mathbf{U})\mathbf{e}_i \mathbf{e}_i^T (\mathbf{A} + \mathbf{R}\mathbf{U})^T \succeq 0. \quad (39)$$

The last inequality is bilinear in t_i and $\lambda_1, \dots, \lambda_m$, and we linearize it by replacing the optimization variable $\boldsymbol{\lambda}$ with a new $\beta_i = t_i \boldsymbol{\lambda}$. Now, the β_i can be assembled into a matrix

$$\boldsymbol{\beta} = [\beta_1 \ \dots \ \beta_n] = \boldsymbol{\lambda} \mathbf{t}, \quad (40)$$

which, as shown above, should have rank 1 and satisfy $\beta_{ij} > 0$, $\mathbf{1}_m^T \boldsymbol{\beta} = \mathbf{t}$. However, the rank-1 constraint for $\boldsymbol{\beta}$ cannot be directly imposed in convex formulations, and we resort to a common technique to indirectly induce low rank in optimal solutions by adding to the cost function the (scaled) nuclear norm $\|\boldsymbol{\beta}\|_N$.

Regarding the second term on the left-hand side of (39), we first note that

$$(\mathbf{A} + \mathbf{R}\mathbf{U})\mathbf{e}_i = [\mathbf{A}\mathbf{e}_i \ \mathbf{R}] \begin{bmatrix} 1 \\ \mathbf{U}\mathbf{e}_i \end{bmatrix} = [\boldsymbol{\alpha}_i \ \mathbf{R}] \begin{bmatrix} 1 \\ \mathbf{v}_i \end{bmatrix}, \quad (41)$$

where $\boldsymbol{\alpha}_i$ and \mathbf{v}_i denote the i -th columns of matrices \mathbf{A} and \mathbf{U} , respectively. Now, consider the following variable, obtained from the stacked rotation vectors that make up \mathbf{U} ,

$$\mathbf{W} = \begin{bmatrix} 1 \\ \text{vec}(\mathbf{U}^T) \end{bmatrix} \begin{bmatrix} 1 & \text{vec}(\mathbf{U}^T)^T \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{u}_1^T & \dots & \mathbf{u}_m^T \\ \mathbf{u}_1 & \underbrace{\mathbf{u}_1 \mathbf{u}_1^T}_{\mathbf{W}_{11}} & & \\ \vdots & & \ddots & \\ \mathbf{u}_m & & & \underbrace{\mathbf{u}_m \mathbf{u}_m^T}_{\mathbf{W}_{mm}} \end{bmatrix}. \quad (42)$$

Further, let I_i denote the set of row indices that extracts the elements of $[1 \ \mathbf{v}_i^T]^T$ in (41) from the first column of \mathbf{W} . Then, the dyad below is readily obtained by selecting the submatrix formed from the I_i rows and I_i columns of \mathbf{W}

$$\mathbf{W}_{I_i I_i} = \begin{bmatrix} 1 \\ \mathbf{v}_i \end{bmatrix} \begin{bmatrix} 1 & \mathbf{v}_i^T \end{bmatrix}, \quad (43)$$

and this carries over to (39) through (41), which can therefore be written in terms of submatrix $\mathbf{W}_{I_i I_i}$. The positive semidefinite matrix \mathbf{W} will replace \mathbf{U} as an optimization variable, retaining the constraints along the diagonal blocks in (42), namely, $\text{tr}(\mathbf{W}_{ii}) = 1$. Finally, we obtain the full convex relaxation of (37) by combining all the above elements and dropping the rank-1 constraint for \mathbf{W} that is implied by (42)

$$\begin{aligned}
& \text{minimize} \quad \mathbf{t}\mathbf{1}_n + \mu\|\boldsymbol{\beta}\|_N \\
& \mathbf{W}, \boldsymbol{\beta}, \mathbf{t} \\
& \text{subject to} \quad \text{diag}(\boldsymbol{\beta}_i) + t_i \sigma \mathbf{1}_m \mathbf{1}_m^T \succeq \begin{bmatrix} \boldsymbol{\alpha}_i & \mathbf{R} \end{bmatrix} \mathbf{W}_{I_i I_i} \begin{bmatrix} \boldsymbol{\alpha}_i^T \\ \mathbf{R} \end{bmatrix} \\
& \mathbf{W} \succeq 0, \quad w_{11} = 1, \quad \text{tr}(\mathbf{W}_{ii}) = 1, \quad \beta_{ij} > 0, \quad \mathbf{1}_m^T \boldsymbol{\beta} = \mathbf{t}.
\end{aligned} \tag{44}$$

This reference formulation for SL- ℓ_1 in multiple dimensions is denoted by SL- ℓ_1 MD.

3) *Simplified non-iterative formulation (SL- ℓ_1 SD)*: Our simulation results suggest that in most scenarios the accuracy of the solution obtained from (44) is nearly identical to that of a simplified formulation where a single epigraph variable, t , is used. Referring to (37), we now minimize $\text{tr}(t\mathbf{I}_n)$ or, equivalently, t , and replace the first constraint for all $i = 1, \dots, n$ with the single matrix inequality $(\mathbf{A} + \mathbf{R}\mathbf{U})^T \boldsymbol{\Xi}(\mathbf{A} + \mathbf{R}\mathbf{U}) \preceq t\mathbf{I}_n$. Applying Schur complements as in (38)–(39) yields

$$t(\mathbf{A} + \sigma \mathbf{1}_m \mathbf{1}_m^T) - \begin{bmatrix} \mathbf{A} & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n \\ \mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{U}^T \end{bmatrix} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{R} \end{bmatrix} \succeq 0, \tag{45}$$

and again we replace variable $\boldsymbol{\lambda}$ with $\boldsymbol{\beta} = t\boldsymbol{\lambda}$ such that $\beta_i > 0$, $\mathbf{1}_m^T \boldsymbol{\beta} = t$. Now, however, there is no need to assemble a matrix as in (40) and to include its nuclear norm as a penalization term in the cost function. Finally, to obtain a convex relaxation we replace \mathbf{U} with the new variable

$$\mathbf{W} = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{U}^T \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & \mathbf{U}^T \\ \underbrace{\mathbf{W}_{11}}_{\mathbf{W}_{11}} & \mathbf{U}^T \mathbf{U} \end{bmatrix}, \tag{46}$$

and drop the rank- n constraint on \mathbf{W} that follows from (46). The simplified SDP formulation for SL- ℓ_1 in multiple dimensions, denoted by SL- ℓ_1 SD, is given by

$$\begin{aligned}
& \text{minimize} \quad t \\
& \mathbf{W}, \boldsymbol{\beta}, t \\
& \text{subject to} \quad \text{diag}(\boldsymbol{\beta}) + t\sigma \mathbf{1}_m \mathbf{1}_m^T \succeq \begin{bmatrix} \mathbf{A} & \mathbf{R} \end{bmatrix} \mathbf{W} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{R} \end{bmatrix} \\
& \mathbf{W} \succeq 0, \quad \mathbf{W}_{11} = \mathbf{I}_n, \quad w_{ii} = 1, \quad \beta_i > 0, \quad \mathbf{1}_m^T \boldsymbol{\beta} = t.
\end{aligned} \tag{47}$$

Note that the optimization variables \mathbf{W} and β in (44) have size $(mn + 1) \times (mn + 1)$ and $m \times n$, respectively, whereas the corresponding sizes in (47) are only $(m + n) \times (m + n)$ and $m \times 1$. For ambient dimension $n = 2$ or 3 and for $m \approx 5$ anchors used in our simulations problem (47) has considerably fewer variables than (44), and the gap increases as m and n grow.

Given the configuration for variable \mathbf{W} in both non-iterative formulations of SL- ℓ_1 (42), (46), the required elements of the rotation vectors that make up \mathbf{U} can be obtained from the rightmost (block) column of \mathbf{W} or by factorizing submatrices along the block diagonal. The former approach is usually more accurate [13].

V. NUMERICAL RESULTS

In this section the tightness and the accuracy of our source localization algorithms are tested in 2D and 3D scenarios, and under various noise assumptions. Results are benchmarked against another relaxation-based method proposed in [8], denoted below as SDR (as in [4]), and with the Squared Range LS (SR-LS) approach of [4]. While the latter does not resort to relaxation, but rather directly optimizes the source coordinates using an iterative root-finding procedure, we take its performance as representative of the current state of the art in optimization-based source localization.

In each reported simulation we performed M Monte Carlo runs, where in each run the source and anchor locations were randomly generated from a uniform distribution over a square or cube whose sides are $[-10, 10]$. The observed ranges, corrupted by i.i.d. noise, were generated as described in Section II under appropriate noise probability densities. The tables list Root Mean-Square Errors (RMSE), computed as $\sqrt{\frac{1}{M} \sum_{i=1}^M \|x_i - \hat{x}_i\|^2}$, where x_i and \hat{x}_i denote the actual and estimated source positions in the i -th Monte Carlo run, respectively. For ease of reference, the best result among all algorithms tested for any given setup will often be shown in boldface.

Convexity and tightness of SLCP: In this example we characterize the accuracy of the convex relaxation used in SLCP and compare its performance to that of the SDR algorithm of [8]. Range measurements to a variable number of randomly placed anchors were generated as indicated above over $M = 1000$ Monte Carlo runs, and corrupted by white Gaussian noise.

First, we estimate how often the constraint set \mathcal{S} (8), which appears in our formulation of the source localization problem prior to relaxation (7), is convex along its “upper right” boundary where the optimal solution lies. As discussed in Section III-A, when this property holds the relaxed solution Φ obtained by SLCP (6) will have rank 1 and can be factorized to yield the optimal point for the non-relaxed problem (7) on the boundary of \mathcal{S} . We empirically assess convexity of \mathcal{S} by tracing the boundary of the

(partially hypothesized) convex hull \mathcal{T} (11) and searching for line segments that delimit regions where the boundaries of \mathcal{S} and \mathcal{T} depart due to local concavity of \mathcal{S} . Specifically, we solve the support hyperplane problem (12) for a grid of angles $0 \leq \beta \leq \frac{\pi}{2}$ and detect the presence of a line segment when the distance between the intersection points $(u(\beta), v(\beta))$ for two consecutive angles β exceeds a threshold. For a noise standard deviation $\sigma_{\text{gaussian}} = 10^{-2}$, \mathcal{S} passed the convexity test in 80% of runs for three anchors. The percentage increased to 84% for five anchors, in line with our reasoning in Section III-A that \mathcal{S} is more likely to be convex as the number of anchors increases.

Next, we compare the RMSEs of SDR and SLCP. As in [5] we provide results for all Monte Carlo runs (denoted by SDR, SLCP) and also for so-called *tight runs* (denoted by SDRt, SLCPt) where the solution for the relaxed localization problem is close to having rank 1, as desired for subsequent factorization to obtain the actual source coordinates. We consider a solution matrix to be tight when the ratio between its first and second eigenvalues is at least 10^2 . Table III lists the RMSEs and the number of tight runs (N_{SDR} , N_{SLCP}) over 1000 trials for five anchors and Gaussian noise standard deviations of 1, 10^{-1} , 10^{-2} , and 10^{-3} . SLCP is clearly superior over the full set of trials, but the gap to SDR closes in the subset of tight runs, indicating that the advantage is mostly due to a much higher probability of its solution having near rank 1. Even for the highest noise power, where the number of tight runs in both algorithms is comparable, the ratio of first to second eigenvalues is usually higher in SLCP, leading to lower RMSE.

TABLE III: Source localization accuracy for relaxation-based methods (RMSEs listed for total and tight runs).

σ_{gaussian}	N_{SDR}	N_{SLCP}	SDR [8]	SDRt [8]	SLCP	SLCPt
10^{-3}	490	921	0.0045	0.0014	0.0020	0.0015
10^{-2}	444	815	0.0162	0.0107	0.0112	0.0108
10^{-1}	478	527	0.1503	0.0960	0.1207	0.0959
1	538	526	1.6070	1.1885	1.2169	1.1885

Under the same simulation setup as above, but using only three anchors, we test the alternative search-based method described in Section III-B to obtain the vector of rotation factors θ from the relaxed solution matrix of SLCP, Φ . Improvements in total RMSE are under 1% for all noise variances using 2×10^5 grid points on the interval $[0, 2\pi)$ to evaluate (17). Foremost, this suggests that rank-1 factorization by SVD, which we adopt as our technique of choice to efficiently extract rotation factors, yields results that are indeed very close to the best possible strategy for finding θ .

Localization in 2D and 3D under Gaussian noise: In the remaining simulations our algorithms are benchmarked against SR-LS [4], whose global performance exceeds that of SDR [8] because it directly optimizes over source locations and does not experience degradations related to tightness of the solutions in the same way that SDR does.

Measurements for the 2D case were generated as in the convexity/tightness assessment, using five anchors and $M = 200$ Monte Carlo runs. Table IV lists the RMSEs for SR-LS, SLCP (6) and its multidimensional counterpart SLNN (27), and also the algorithms for Laplacian noise, namely, the complex formulation of $\text{SL-}\ell_1$ in [3], and its multidimensional counterpart $\text{SL-}\ell_1$ AD (34), (36). All

TABLE IV: 2D source localization under Gaussian noise (5 anchors, 200 Monte Carlo runs). RMSEs are given for complex (SLCP, $\text{SL-}\ell_1$) and real (SLNN, $\text{SL-}\ell_1$ AD) formulations.

σ_{gaussian}	SR-LS [4]	SLCP	$\text{SL-}\ell_1$ [3]	SLNN	$\text{SL-}\ell_1$ AD
10^{-3}	0.0032	0.0023	0.0014	0.0023	0.0057
10^{-2}	0.0138	0.0109	0.0136	0.0113	0.0133
10^{-1}	0.1406	0.1037	0.1118	0.1097	0.1249
1	1.4947	1.3249	1.4536	1.3580	1.4593

algorithms outperform SR-LS, which squares measurements ($p = 2$, $q = 2$ in (1)) and thus becomes more sensitive to the presence of (Gaussian) noise in range measurements. The fact that SLCP/SLNN achieve the best results is not surprising, as these algorithms actually maximize a Gaussian likelihood function. Interestingly, SLCP attains slightly lower errors than SLNN, even though the same cost function and similar steps are used in the derivation of both algorithms. However, the impact of relaxing the rank constraint in the optimization variable to obtain a semidefinite program is not necessarily the same, which could explain the observed differences in performance. Similar comments apply to $\text{SL-}\ell_1$ and $\text{SL-}\ell_1$ AD, although the algorithmic differences between the complex and real formulations are larger than for SLCP/SLNN.

Results for 3D source localization, for which the complex formulations cannot be used, are given in Table V. Again, both SLNN and $\text{SL-}\ell_1$ outperform SR-LS, the former having lower RMSE as its cost function is matched to the noise statistics. Source localization with the same number of anchors is a less constrained problem in 3D than it is in 2D, resulting in higher RMSEs. Regarding the three variants of $\text{SL-}\ell_1$ (see Section IV-A), there is no well-defined trend on their relative performance. Note also how the RMSE of the non-adaptive simplified formulation $\text{SL-}\ell_1$ SD (47) is quite close to that of the general

TABLE V: 3D source localization under Gaussian noise (5 anchors, 200 Monte Carlo runs).

σ_{gaussian}	SR-LS [4]	SLNN	SL- ℓ_1		
			AD	MD	SD
10^{-3}	0.0040	0.0036	0.0038	0.0038	0.0040
10^{-2}	0.0295	0.0274	0.0285	0.0290	0.0292
10^{-1}	0.2612	0.2290	0.2376	0.2401	0.2390
1	3.3279	2.7431	2.9492	2.8748	2.8801

formulation SL- ℓ_1 MD (44), even outperforming it, on average, for one of the noise powers.

Localization in 2D and 3D in the presence of outliers: The same setup for Gaussian noise is adopted here, except that ranges are contaminated either by Laplacian noise, or by what we designate as *selective Gaussian noise*. Range measurements for the latter are created as $r_i = \|x - a_i\| + w_i + |\epsilon|$, where w_i is a Gaussian noise term with $\sigma_{\text{gaussian}} = 0.04$ that is present in all observations and ϵ is also a Gaussian disturbance, but with higher standard deviation $\sigma_{\text{outlier}} \in [0.3, 1.5]$, that contaminates only one measured range (i.e., $\epsilon = 0$ for all other observations). This statistical model is less tractable than the Laplacian noise model, but we include it in some of our simulations as it more realistically reflects how outliers occur in real ranging systems.

Tables VI and VII list RMSEs for 2D and 3D source localization under both outlier generation models. When compared with Tables IV and V for the Gaussian case, the most striking difference is that the variants of SL- ℓ_1 , designed for Laplacian noise, now outperform both SR-LS and SLCP/SLNN. As in the 2D Gaussian case, the complex formulation SL- ℓ_1 attains lower errors than the real formulation SL- ℓ_1 AD. In 3D (Table VII) the real variants of SL- ℓ_1 developed in Section IV-A provide significant RMSE reductions, on the order of 10%, over those of the Gaussian algorithm SLNN. Interestingly, this conclusion still holds for the selective Gaussian case, whose outlier generation model is not matched to the Laplacian assumption underlying SL- ℓ_1 . Similarly to the Gaussian case of Table V, the three variants of SL- ℓ_1 exhibit similar performance, here with a slight advantage of the alternating direction algorithm SL- ℓ_1 AD.

VI. CONCLUSION

We have proposed SLNN as an extension to 3 and higher dimensions of the ML-based source localization approach developed in [3] by formulating it as an optimization problem using nuclear norms and SDR. Similarly, we also extended to higher dimensions the 2D SL- ℓ_1 localization algorithm for Laplacian

TABLE VI: 2D source localization in the presence of outlier range measurements (5 anchors, 200 Monte Carlo runs).

(a) Laplacian noise

$\sigma_{\text{laplacian}}$	SR-LS [4]	SLCP	SL- ℓ_1 [3]	SLNN	SL- ℓ_1 AD
0.2	1.1398	1.0839	0.9240	1.0770	1.0687
0.4	1.9031	1.8585	1.4546	1.8908	1.7996
0.8	3.1543	3.1143	3.0344	3.0812	3.0798

(b) Selective Gaussian noise

σ_{outlier}	SR-LS [4]	SLCP	SL- ℓ_1 [3]	SLNN	SL- ℓ_1 AD
0.5	0.2983	0.2448	0.1849	0.2556	0.2337
1.0	0.4662	0.4561	0.2508	0.4516	0.3714
1.5	1.2419	1.1640	1.0542	1.2389	1.2157

noise developed in [9]. Our simulation results show that the proposed algorithms provide very accurate results compared to other optimization-based localization methods that operate on range measurements, although their performance in 2D is not quite as good as that of the complex formulations developed in [3], [9]. In 3D scenarios with Gaussian noise SLNN delivered solutions that were about 5% more accurate than those of SL- ℓ_1 , whereas in the presence of outlier range measurements the situation was reversed and SL- ℓ_1 proved to be about 5–10% more accurate under either Laplacian or selective Gaussian models. We developed both iterative (SL- ℓ_1 AD) and non-iterative (SL- ℓ_1 MD/SD) 3D extensions of SL- ℓ_1 , which exhibited comparable performance, with a slight advantage of SL- ℓ_1 AD. Complexity considerations (e.g., computational load, maximum admissible problem size) will then play an important role when selecting one of those algorithms for a particular application.

We have carried out an analysis of the geometry of our 2D formulation for ML localization under Gaussian noise (SLCP), and found that the high probability that a certain portion of the (outer) border of its constraint set is convex justifies the observed strong tightness of our relaxation. Our simulation results for random anchor configurations indicate that another well-known SDR relaxation for the same problem has a significantly higher chance of yielding optimal solutions that do not have the necessary properties (unit rank) to accurately recover source positions. Regarding the extraction of spatial coordinates from the positive semidefinite matrix computed by SLCP, we examined a search-based alternative to standard rank-1 factorization using the SVD. This strategy is feasible for the practically important case of range-

TABLE VII: 3D source localization in the presence of outlier range measurements (5 anchors, 200 Monte Carlo runs).

(a) Laplacian noise

$\sigma_{\text{laplacian}}$	SR-LS [4]	SLNN	SL- ℓ_1		
			AD	MD	SD
0.25	1.3619	1.3577	1.2113	1.2097	1.1776
0.5	2.6514	2.5719	2.3236	2.4704	2.3651
0.75	3.5968	3.5070	3.1265	3.2173	3.1987

(b) Selective Gaussian noise

σ_{outlier}	SR-LS [4]	SLNN	SL- ℓ_1		
			AD	MD	SD
0.3	0.3930	0.3237	0.3033	0.3035	0.3033
0.6	0.9756	0.9154	0.8786	0.9084	0.9001
0.9	1.5934	1.5502	1.2755	1.2857	1.3022

based localization using three anchors, but was found to yield only minor improvements relative to the SVD-based factorization.

APPENDIX A

PROOF OF LEMMA 1

Our proof of Lemma 1 relies on a result, interesting in its own right, that characterizes the convex hull of the set of 3×3 rank-1 matrices built from complex vectors with unit-magnitude components.

Lemma 2. *Let*

$$\mathcal{A} = \{ \boldsymbol{\theta} \boldsymbol{\theta}^H : \boldsymbol{\theta} \in \mathbb{C}^3, |\theta_i| = 1 \}, \quad (48)$$

$$\mathcal{B} = \{ \boldsymbol{\Phi} \in \mathbb{C}^{3 \times 3} : \boldsymbol{\Phi} \succeq 0, \phi_{ii} = 1 \}. \quad (49)$$

then $\mathcal{B} = \text{co}(\mathcal{A})$.

Proof: $\text{co}(\mathcal{A}) \subset \mathcal{B}$ is straightforward since \mathcal{B} is convex and $\mathcal{A} \subset \mathcal{B}$. For the reverse direction $\text{co}(\mathcal{A}) \supset \mathcal{B}$ our goal is to find, for every $\boldsymbol{\Phi} \in \mathcal{B}$, matrices $\boldsymbol{\Theta}_i \in \mathcal{A}$ and nonzero scalars $\lambda_i \geq 0$, with $\sum_i \lambda_i = 1$, such that $\boldsymbol{\Phi} = \sum_i \lambda_i \boldsymbol{\Theta}_i$.

Note that both \mathcal{A} and \mathcal{B} are invariant under the (unitary) similarity operation

$$\mathbf{M} \rightarrow \mathbf{PMP}^H, \quad (50)$$

where \mathbf{P} is the product of a permutation and a diagonal unitary matrix. In other words, we can simultaneously permute rows and columns and multiply the i -th row and i -th column by a unit-magnitude complex number. Thus, we can assume without loss of generality that Φ is of the form

$$\Phi = \begin{bmatrix} 1 & a & b \\ a & 1 & z^* \\ b & z & 1 \end{bmatrix}, \quad 0 \leq a \leq b, \quad z \in \mathbb{C}. \quad (51)$$

Since $\Phi \succeq 0$, we must have $a \leq 1$, $b \leq 1$, $|z| \leq 1$ and

$$0 \leq |\Phi| = 1 - a^2 - b^2 - |z|^2 + 2ab \operatorname{Re}\{z\}, \quad (52)$$

which, for $z = x + jy$, reads

$$(x - ab)^2 + y^2 \leq (1 - a^2)(1 - b^2). \quad (53)$$

For fixed a, b this inequality describes a circle (with interior) in the (x, y) plane, centered on $(ab, 0)$. Since any point in the interior of a circle can be written as a convex combination of two points on its boundary, we can assume that we have equality in (53). Thus, from now on we assume

$$z = ab + \sqrt{(1 - a^2)(1 - b^2)}e^{j\varphi}. \quad (54)$$

We now complete the proof by expressing such Φ as a convex combination of two matrices from \mathcal{A} . For given $0 \leq a \leq b \leq 1$ and $\varphi \in [0, 2\pi[$ we want to find $\alpha, \beta, \gamma, \delta \in [0, 2\pi[$, and $0 \leq \lambda \leq 1$ such that

$$\Phi = \begin{bmatrix} 1 & a & b \\ a & 1 & z^* \\ b & z & 1 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ e^{j\alpha} \\ e^{j\beta} \end{bmatrix} \begin{bmatrix} 1 & e^{-j\alpha} & e^{-j\beta} \end{bmatrix} + (1 - \lambda) \begin{bmatrix} 1 \\ e^{j\gamma} \\ e^{j\delta} \end{bmatrix} \begin{bmatrix} 1 & e^{-j\gamma} & e^{-j\delta} \end{bmatrix}. \quad (55)$$

We thus have

$$a = \lambda e^{j\alpha} + (1 - \lambda)e^{j\gamma}, \quad b = \lambda e^{j\beta} + (1 - \lambda)e^{j\delta}, \quad (56)$$

$$z = \lambda e^{j(\beta - \alpha)} + (1 - \lambda)e^{j(\delta - \gamma)}. \quad (57)$$

From the first two relations we get

$$e^{j\gamma} = \frac{a - \lambda e^{j\alpha}}{1 - \lambda}, \quad e^{j\delta} = \frac{b - \lambda e^{j\beta}}{1 - \lambda}, \quad (58)$$



Fig. 3: Illustration of geometrical Lemma 3.

and replacing these in the third relation yields, after simple manipulations,

$$z = ab + \frac{\lambda}{1-\lambda}(e^{-j\alpha} - a)(e^{j\beta} - b). \quad (59)$$

Before proceeding, we state and prove a useful lemma from elementary geometry:

Lemma 3. Referring to Figure 3a, if A is a point inside a unit circle whose distance to the center is a , RS is any line through A , and PQ is a diameter through A , then

$$AR \cdot AS = AP \cdot AQ = (1-a)(1+a) = 1-a^2. \quad (60)$$

Proof: Triangles APR and AQS , depicted in Figure 3b, are similar, hence

$$\frac{AP}{AS} = \frac{AR}{AQ}. \quad (61)$$

■

We use the lemma above with parameters as depicted in Figure 4. From $A = \lambda R + (1-\lambda)S$ we have $\frac{AR}{AS} = \frac{1-\lambda}{\lambda}$, and by Lemma 3 $AR \cdot AS = 1 - a^2$, hence

$$AR = \sqrt{\frac{1-\lambda}{\lambda}(1-a^2)}, \quad e^{j\alpha} = a + \sqrt{\frac{1-\lambda}{\lambda}(1-a^2)}e^{j\varphi_1}. \quad (62)$$

Similarly, with $A = b$, $R = e^{j\beta}$, $S = e^{j\delta}$, and φ_2 instead of φ_1 , we have

$$e^{j\beta} = b + \sqrt{\frac{1-\lambda}{\lambda}(1-b^2)}e^{j\varphi_2}. \quad (63)$$

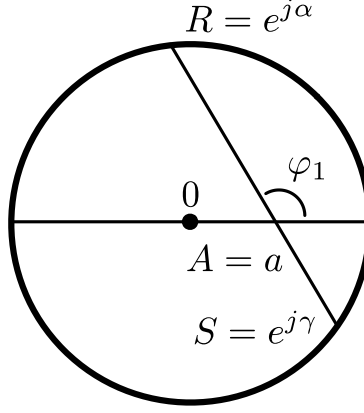


Fig. 4: Application of Lemma 3 to a convex combination on the unit circle.

Substituting (62), (63) back in (59) yields

$$z = ab + \sqrt{(1-a^2)(1-b^2)}e^{j(\varphi_2-\varphi_1)}, \quad (64)$$

which has the same form as (54), obtained from the positive semidefinite condition for matrix Φ in (51).

We now argue that letting angle α go from 0 to 2π is equivalent to letting φ_1 cover an interval of length 2π as well (Figure 4). Fixing φ_1 , and consequently α , the two relations in (56), together with an arbitrary requirement that $\text{Im}\{e^{j\beta}\} \geq 0$, fix the values² of β , γ , δ , λ , and, in particular, of φ_2 . Thus, $\varphi_2 = f(\varphi_1)$ is a continuous function of φ_1 .

When $\varphi_1 = 0$, φ_2 has a certain value, say, $\varepsilon_0 \in [0, \pi]$ (it can be computed, but is not needed in this proof). For $\varphi_1 = \pi$ it is straightforward to see that $\varphi_2 = \pi - \varepsilon_0$, and for $\varphi_1 = 2\pi$ it is again ε_0 . In particular the continuous function $\varphi_2 - \varphi_1$ takes values from $\varepsilon_0 - 0 = \varepsilon_0$ to $\varepsilon_0 - 2\pi$, i.e., modulo 2π it takes all values in $[0, 2\pi[$. Thus, for any given angle φ in (54), let φ_1 be such that $f(\varphi_1) - \varphi_1 = \varphi$, modulo 2π . Then, the corresponding α , β , γ , δ , and λ , as explained above, give the desired decomposition (55). ■

We now proceed and prove Lemma 1 under the assumption of Lemma 2, thus tacitly assuming $m = 3$. Note that the proof is valid for arbitrary \mathbf{c} , \mathbf{r} in (8) and (11), i.e., it does not require that the structure for these vectors defined in (4) be taken into account.

²Equivalently, note that fixing φ_1 fully defines the geometrical construction shown in Figure 4, and thus fixes the values of γ and λ . Then, λ fully defines the corresponding construction for $A = b$ if, in addition, $\text{Im}\{e^{j\beta}\} \geq 0$ is specified, and thus fixes the values of β , δ , and φ_2 .

Proof: We rewrite sets \mathcal{S} in (8) and \mathcal{T} in (11) using the notation (48)

$$\mathcal{S} = \{(\mathbf{c}^H \Theta \mathbf{c}, \mathbf{r}^T \Theta \mathbf{r}) : \Theta \in \mathcal{A}\}, \quad (65)$$

$$\mathcal{T} = \{(\mathbf{c}^H \Phi \mathbf{c}, \mathbf{r}^T \Phi \mathbf{r}) : \Phi \in \text{co}(\mathcal{A})\}. \quad (66)$$

Obviously $\mathcal{S} \subset \mathcal{T}$. Now let $\alpha \in [0, \frac{\pi}{2}]$ and define

$$(u_1, v_1) = \arg \max_{(u,v) \in \mathcal{T}} \langle (\cos \alpha, \sin \alpha), (u, v) \rangle. \quad (67)$$

We wish to show that

$$\langle (\cos \alpha, \sin \alpha), (u_1, v_1) \rangle = \max_{(u,v) \in \mathcal{S}} \langle (\cos \alpha, \sin \alpha), (u, v) \rangle, \quad (68)$$

so that the inner product over \mathcal{S} attains the same maximum value as over the larger set \mathcal{T} , and the support hyperplanes with normal $(\cos \alpha, \sin \alpha)$ thus coincide for the two sets. It is enough to prove that there exists $(u', v') \in \mathcal{S}$ that attains the left-hand side of (68).

We may write $\Phi_1 \in \text{co}(\mathcal{A})$ which maximizes (67) as

$$\Phi_1 = \sum_i \lambda_i \theta_i \theta_i^H, \quad \lambda_i \geq 0, \quad \sum_i \lambda_i = 1, \quad |\theta_{ik}| = 1, \quad (69)$$

hence

$$\begin{aligned} & \langle (\cos \alpha, \sin \alpha), (u_1, v_1) \rangle \\ &= (\sqrt{\cos \alpha} \mathbf{c})^H \Phi_1 (\underbrace{\sqrt{\cos \alpha} \mathbf{c}}_{\mathbf{p}}) + (\sqrt{\sin \alpha} \mathbf{r})^T \Phi_1 (\underbrace{\sqrt{\sin \alpha} \mathbf{r}}_{\mathbf{q}}) \\ &= \sum_i \lambda_i (\mathbf{p}^H \theta_i \theta_i^H \mathbf{p} + \mathbf{q}^H \theta_i \theta_i^H \mathbf{q}) \\ &= \sum_i \lambda_i (|\mathbf{p}^H \theta_i|^2 + |\mathbf{q}^H \theta_i|^2). \end{aligned} \quad (70)$$

Let i_0 be the index where the last summation attains its maximum value. Then

$$\langle (\cos \alpha, \sin \alpha), (u_1, v_1) \rangle \leq |\mathbf{p}^H \theta_{i_0}|^2 + |\mathbf{q}^H \theta_{i_0}|^2 = \langle (\cos \alpha, \sin \alpha), (\mathbf{c}^H \theta_{i_0} \theta_{i_0}^H \mathbf{c}, \mathbf{r}^T \theta_{i_0} \theta_{i_0}^H \mathbf{r}) \rangle, \quad (71)$$

which completes the proof because the second argument in the inner product is an element of \mathcal{S} . ■

APPENDIX B

ANALYSIS OF SLNN

Solution of the inner subproblem (23): For any optimization problem with differentiable objective and constraint functions for which strong duality holds, any set of primal and dual optimal points must satisfy the KKT conditions [16]. We define the Lagrangian of (23) with dual variable Λ as

$$L(\mathbf{V}, \Lambda) = \text{tr}(\mathbf{C}^T \mathbf{U} \mathbf{V}) + \text{tr}(\Lambda^T (\mathbf{V}^T \mathbf{V} - \mathbf{I}_n)). \quad (72)$$

The first-order KKT conditions are given by

$$\nabla_{\mathbf{V}} L(\mathbf{V}, \mathbf{\Lambda}) = \mathbf{U}^T \mathbf{C} + \mathbf{V}(\mathbf{\Lambda} + \mathbf{\Lambda}^T) = \mathbf{0}, \quad (73)$$

$$\mathbf{V}^T \mathbf{V} = \mathbf{I}_n, \quad (74)$$

where (73) is obtained by setting to zero the gradient³ of (72) with respect to \mathbf{V} , whereas (74) is the original orthogonality constraint in (23).

Premultiplying (73) with \mathbf{V}^T , taking the trace (i.e., taking the inner product with \mathbf{V}), and using (74) yields the optimal value for the cost function

$$\text{tr}(\mathbf{C}^T \mathbf{U} \mathbf{V}) = \text{tr}(\mathbf{V}^T \mathbf{U}^T \mathbf{C}) = -\text{tr}(\mathbf{\Lambda} + \mathbf{\Lambda}^T). \quad (75)$$

But from $\mathbf{U}^T \mathbf{C} = -\mathbf{V}(\mathbf{\Lambda} + \mathbf{\Lambda}^T)$ in (73) we can square both sides to get

$$\mathbf{C}^T \mathbf{U} \mathbf{U}^T \mathbf{C} = (\mathbf{\Lambda} + \mathbf{\Lambda}^T)^2. \quad (76)$$

Hence, among candidate optimal points satisfying the KKT system, the cost function can be made as small as possible by choosing $\mathbf{\Lambda} + \mathbf{\Lambda}^T$ in (75) as a positive semidefinite matrix square root of the left-hand side of (76). Replacing this in (75) gives the final optimal cost

$$\text{tr}(\mathbf{C}^T \mathbf{U} \mathbf{V}) = -\text{tr}((\mathbf{C}^T \mathbf{U} \mathbf{U}^T \mathbf{C})^{\frac{1}{2}}) = -\|\mathbf{C}^T \mathbf{U}\|_N. \quad (77)$$

■

Interestingly, we point out that the more usual Frobenius norm solves the following relaxed version of the inner subproblem (23)

$$\begin{aligned} & \text{minimize} \quad \text{tr}(\mathbf{C}^T \mathbf{U} \mathbf{V}) = \langle \mathbf{V}, \mathbf{U}^T \mathbf{C} \rangle \\ & \quad \mathbf{V} \end{aligned} \quad (78)$$

$$\text{subject to} \quad \text{tr}(\mathbf{V}^T \mathbf{V}) = \|\mathbf{V}\|_F^2 = n,$$

which is easily verified by writing the KKT system based on the Lagrange function $\text{tr}(\mathbf{C}^T \mathbf{U} \mathbf{V}) + \lambda(\text{tr}(\mathbf{V}^T \mathbf{V}) - n)$,

$$\mathbf{U}^T \mathbf{C} + 2\lambda \mathbf{V} = \mathbf{0}, \quad \text{tr}(\mathbf{V}^T \mathbf{V}) = n, \quad (79)$$

whose solution at the minimum is

$$\mathbf{V} = -\sqrt{n} \frac{\mathbf{U}^T \mathbf{C}}{\|\mathbf{U}^T \mathbf{C}\|_F}, \quad \lambda = \frac{\|\mathbf{U}^T \mathbf{C}\|_F}{2\sqrt{n}}, \quad (80)$$

³We use the standard results $\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A}^T \mathbf{X}) = \mathbf{A}$ and $\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X} \mathbf{B} \mathbf{X}^T) = \mathbf{X}(\mathbf{B} + \mathbf{B}^T)$ [17], [19].

with optimal cost $-\sqrt{n}\|\mathbf{U}^T\mathbf{C}\|_F$. The minimum cost within the expanded domain of this relaxed subproblem will at least be as low as that of (23), hence $\|\mathbf{U}^T\mathbf{C}\|_N \leq \sqrt{n}\|\mathbf{U}^T\mathbf{C}\|_F$. On the other hand,

$$\|\mathbf{U}^T\mathbf{C}\|_N = \sqrt{\left(\sum_i \sigma_i\right)^2} \geq \sqrt{\sum_i \sigma_i^2} = \|\mathbf{U}^T\mathbf{C}\|_F, \quad (81)$$

where σ_i denotes the i -th singular value of $\mathbf{U}^T\mathbf{C}$. Combining the two inequalities we have the bounds

$$\|\mathbf{U}^T\mathbf{C}\|_F \leq \|\mathbf{U}^T\mathbf{C}\|_N \leq \sqrt{n}\|\mathbf{U}^T\mathbf{C}\|_F. \quad (82)$$

Proof of equivalence between (26) and (27): We first rewrite (27) replacing the linear matrix inequality with an equivalent Schur complement

$$\begin{aligned} & \text{maximize} \quad 2\text{tr}(\mathbf{Z}) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{W}) \\ & \quad \mathbf{W}, \mathbf{Z} \\ & \text{subject to} \quad \mathbf{W} \succeq 0, \quad W_{ii} = 1 \\ & \quad \mathbf{Z}^2 \preceq \mathbf{C}^T\mathbf{W}\mathbf{C}, \quad \mathbf{Z} \succeq 0. \end{aligned} \quad (83)$$

Let p_1^* and p_2^* be the optimal values of problems (26) and (83), respectively.

Choose a feasible point (\mathbf{Z}, \mathbf{W}) for the second problem, such that $0 \preceq \mathbf{Z}^2 \preceq \mathbf{C}^T\mathbf{W}\mathbf{C}$. This implies⁴ $\mathbf{Z} \preceq (\mathbf{C}^T\mathbf{W}\mathbf{C})^{\frac{1}{2}}$, hence the values of the two objective functions satisfy

$$2\text{tr}(\mathbf{Z}) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{W}) \leq 2\text{tr}((\mathbf{C}^T\mathbf{W}\mathbf{C})^{\frac{1}{2}}) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{W}). \quad (84)$$

In particular, choosing for (\mathbf{Z}, \mathbf{W}) the unique maximizer of (83), inequality (84) asserts that $p_1^* \geq p_2^*$.

For the converse choose a feasible point \mathbf{W} for the first problem and consider the eigendecomposition $\mathbf{C}^T\mathbf{W}\mathbf{C} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$. Now set $\mathbf{Z} = \mathbf{Q}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{Q}^T$, so that $\mathbf{Z}^2 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T = \mathbf{C}^T\mathbf{W}\mathbf{C}$, and (\mathbf{W}, \mathbf{Z}) is therefore feasible for (83). For both problems the value of the cost function is

$$2\text{tr}(\mathbf{\Lambda}^{\frac{1}{2}}) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{W}). \quad (85)$$

In particular, choosing for \mathbf{W} the maximizer of (26) the construction for \mathbf{Z} yields a feasible point (\mathbf{W}, \mathbf{Z}) for (83) where the objective function equals p_1^* . Therefore $p_1^* \leq p_2^*$, and coupling this with the converse inequality above we conclude that $p_1^* = p_2^*$ and the two problems are equivalent. \blacksquare

⁴ $\mathbf{A} \succeq \mathbf{B} \succeq 0 \Rightarrow \mathbf{A}^{\frac{1}{2}} \succeq \mathbf{B}^{\frac{1}{2}} \succeq 0$ [20].

REFERENCES

- [1] W. Dargie and C. Poellabauer, *Fundamentals of Wireless Sensor Networks: Theory and Practice*, ser. Wiley Series on Wireless Communications and Mobile Computing. John Wiley & Sons, 2010.
- [2] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, “Range-free localization schemes for large scale sensor networks,” in *Proc. ACM MobiCom’03*, San Diego, CA, USA, September 2003.
- [3] P. Oğuz-Ekim, J. Gomes, J. Xavier, and P. Oliveira, “A convex relaxation for approximate maximum-likelihood 2D source localization from range measurements,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP’10)*, Dallas, Texas, USA, March 2010.
- [4] A. Beck, P. Stoica, and J. Li, “Exact and approximate solutions of source localization problems,” *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1770–1778, May 2008.
- [5] A. Beck, M. Teboulle, and Z. Chikishev, “Iterative minimization schemes for solving the single source localization problem,” *SIAM J. on Optimization*, vol. 19, no. 3, pp. 1397–1416, 2008.
- [6] E. Xu, Z. Ding, and S. Dasgupta, “Wireless source localization based on time of arrival measurement,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP’10)*, Dallas, Texas, USA, March 2010.
- [7] Y. Chan, H. Hang, and P. Ching, “Exact and approximate maximum likelihood localization algorithms,” *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 10–16, January 2006.
- [8] K. W. Cheung, W. K. Ma, and H. C. So, “Accurate approximation algorithm for TOA-based maximum likelihood mobile location using semidefinite programming,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP’04)*, Montreal, Canada, May 2004.
- [9] P. Oğuz-Ekim, J. Gomes, J. Xavier, and P. Oliveira, “Robust localization of nodes and time-recursive tracking in sensor networks using noisy range measurements,” *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1770–1778, May 2011.
- [10] A. O. Hero and D. Blatt, “Sensor network source localization via projection onto convex sets (POCS),” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP’05)*, vol. 3, Philadelphia, USA, March 2005, pp. 689–692.
- [11] F. K. W. Chan and H. C. So, “Accurate distributed range-based positioning algorithm for wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4100–4105, 2009.
- [12] A. G. Dimakis, S. Kar, J. M. F. Moura, M. Rabbat, and A. Scaglione, “Gossip algorithms for distributed signal processing,” *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.
- [13] P. Biswas and Y. Ye, “Semidefinite programming for ad hoc wireless sensor network localization,” in *Proc. Int. Conf. Inf. Process. Sensor Netw. (IPSN’04)*, Berkeley, California, USA, April 2004, pp. 46–54.
- [14] K. Yang, G. Wang, and Z.-Q. Luo, “Efficient convex relaxation methods for robust target localization by a sensor network using time differences of arrivals,” *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2775–2784, 2010.
- [15] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins University Press, 1996.
- [16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [17] K. B. Petersen and M. S. Pedersen, “The matrix cookbook,” October 2008, version 20081110. [Online]. Available: <http://www2.imm.dtu.dk/pubdb/p.php?3274>
- [18] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. SIAM, 2001.
- [19] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 3rd ed., ser. Wiley Series in Probability and Statistics. John Wiley & Sons, 2007.
- [20] R. A. Horn and C. A. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.